

Magyar generatív történeti mondattan

Kutatási beszámoló

1. Az ómagyar korpusz

A projektum két feladatot tűzött maga elé és oldott meg. Egyrészt létrehoztuk az ómagyar és korai középmagyar kéziratos szövegek elektronikus adatbázisát. A szövegeket digitalizáltuk, 11 kódex szövegének normalizált – könnyen olvasható és nyomtatható – változatát is elkészítettük, négy kódexet pedig morfológiailag elemeztünk és annotáltunk. A szövegek, beleértve a normalizált és az annotált változatokat is, kereshetők. Mind az adatbázis, mind a kereshető szabadon elérhető, bárki számára használható. Ezzel nagy mértékben megkönnyítettük a magyar nyelvtörténet és különösen az ómagyar kutatását, és lehetővé tettük a mai nemzetközi nyelvtörténeti kutatásokra jellemző statisztikai módszerek alkalmazását. Az adatbázis nemzetközi jelentőséggel is bír. Hasonló nyelvtörténeti adatbázisa – egy kivételtől eltekintve – csak néhány indo-európai nyelvnek van. A különféle nyelvcsaládokba tartozó nyelvek változásainak összehasonlítása alapvető fontosságú a nyelvi változások általános elmélete számára. 2014 nyarán műhelykonferenciát rendezünk az MTA Nyelvtudományi Intézetében a különféle nyelvek történeti adatbázisainak szenderdizálásáról, bizonyos fokú egységesítéséről.

A feldolgozott anyag 47 kódexet, 24 rövidebb szövegemléket és 244 misszilit (elküldött levelet) foglal magába, vagyis mindösszesen 2,2 millió szövegszót.

A korpuszépítés első lépése a valamilyen elektronikus szöveges formátumban már meglévő nyelvtörténeti anyagok összegyűjtése volt. A különböző forrásokból (kiadóktól, nyelvtörténészekről) származó, változatos fontkészleteket használó dokumentumokat egységes, UTF-8 kódolású, sztenderd Unicode-karaktereket tartalmazó sima szövegfájlokká alakítottuk.

Másik forrásunk a Számítógépes Nyelvtörténeti Adattár volt, amelyben több ómagyar kódex ábécérendes adattára elérhető (Guary, Festetics, Jókai). Az adattárban a kódex címszavai (a szövegszavak tövei mai magyar átírásban) ábécérendbe rendezve szerepelnek. A hozzájuk tartozó betűhű szövegszavakat a lelőhely (lapszám, sorszám) megjelölésével közlik, mellettük számokkal rögzítették az adatra vonatkozó helyesírás-történeti, szótörténeti, hangtani, szófajtani, jelentéstani és alaktani tudnivalókat. A szövegben sokszor előforduló szavakat függelékben különítették el. Ezeket a lelőhely alapján visszahelyeztünk az eredeti kódexbeli helyükre. Az egyes szövegszavak soron belüli sorrendjét nem közlik, ezért a sorbarendezt is elvégeztük. Ezután a többféle fontkészletet alkalmazó táblázatot UTF-8

kódolású sima szöveggé alakítottuk, majd ebből állítottuk vissza a kódexek eredeti betűhű szövegét. Az egyes szövegszavakhoz tartozó morfológiai elemzést az általunk használt morfológiai elemző kimeneti formátumára alakítottuk, továbbá a mai magyar tövek és az elemzés alapján rekonstruáltuk a normalizált szóalakot. Ennek a konvertálási munkafolyamatnak a végén megkaptuk az adott kódex szavainak betűhű és normalizált alakját, valamint a hozzájuk tartozó egyértelmű morfológiai elemzést.

Az ómagyar szövegek nagy részének azonban nem volt elektronikusan elérhető szöveges változata, így ezeket a számítógép által olvasható és feldolgozható formára kellett hoznunk. Ez a rövidebb szövegek esetében általában begépeléssel, a hosszabbak esetében szkenneléssel, optikai karakterfelismerő (OCR) program alkalmazásával és kézi ellenőrzéssel történt.

A projektum első felében kidolgoztuk és felépítettük az adatbázisépítés munkamenetét. A szövegek összegyűjtésével és egységes elektronikus formára hozásával párhuzamosan zajlottak a szövegfeldolgozás további lépései is, melyek során a betűhű változat mellett előállítottuk néhány szövegemlék normalizált és morfológiailag elemzett és egyértelműsített változatát is. Ezek alapján a szövegfeldolgozási munkamenet a következő:

1. kiadott kódex szkennelve
→ OCR
2. nyers OCR-kimenet
→ kézi javítás
3. betűhű elektronikus forma
→ kézi normalizálás
4. normalizált forma
→ automatikus morfológiai elemzés
5. szótövesített és morfológiailag elemzett forma
→ kézi egyértelműsítés
6. egyértelműsített korpusz

A különböző szintű szövegek előállítása az automatikus gépi támogatás mellett igen sok kézi kódolási munkát igényelt. A kódolók munkáját segítő és egységesítő kidolgoztunk egy kódolási szabályzatot, melyben rögzítettük a normalizálás alapelveit, a szöveghelyek (lókuszok) jelölésének módját, a mondatrabontás és a szkriptor javításai kódolásának szabályait.

Ahhoz, hogy a korpuszban a nyelvi jelenségek kereshetők legyenek, vagyis az adatbázis használható segédeszköze legyen az elméleti nyelvészeti és nyelvtörténeti kutatásoknak, a releváns információkat elektronikusan előhívható és interpretálható módon kell tárolni. A kifinomult, nyelvészetileg releváns lekérdezések sok esetben különféle nyelvi szinteken megjelenő információra hivatkoznak. Hogy ezek mind elérhetőek legyenek, adatbázisunk párhuzamosan tartalmazza az egyes szövegfeldolgozottsági szinteknek megfelelő nyelvi adatokat. Vagyis az egyes szövegszavakhoz a következő adatok tartoznak:

- betűhű forma = (3)
- normalizált alak = (4)
- szótó – egyértelműsített morfológiai elemzés (6) alapján
- elemzés – egyértelműsített morfológiai elemzés (6) alapján
-

Az adatbázishoz készült egy lekérdező felület, amely nyilvánosan elérhető a <http://rmk.nytud.hu> címen (angol verziója: <http://ohc.nytud.hu>). A lekérdező lényege, hogy bármely szinten meg lehet fogalmazni a kérdésünket. Ha pl. bizonyos speciális karaktereket tartalmazó szavakat akarunk kilistázni, akkor a betűhű szinthez kell fordulnunk. Ha arra vagyunk kíváncsiak, hogy milyen szavak szerepelhetnek egy igealak és a hozzá tartozó igeekötő között, akkor ehhez már tudnunk kell a szavakra vonatkozó szófaji információkat is, vagyis az elemzett és egyértelműsített (6)-os szinten tudjuk a kérdésünket megfogalmazni. A lekérdező felületen keresztül az összes nyelvelmélet betűhű szövege elérhető, valamint bizonyos szövegek normalizált és morfológiaileg elemzett változata is.

A projektum eredményeit egy weboldalon adjuk közre (<http://omagyarkorpusz.nytud.hu>). A projektum és a korpuszépítés részletes magyar és angol nyelvű leírása, a résztvevő munkatársak publikációi és az összes feldolgozott ó- és középmagyar anyag elérhető az oldalon. Minden szövegelemlethez tartozik egy rövid leírás, amelyben ismertetjük, hogy milyen forrásból dolgoztunk, hogy milyen feldolgozottsági szinten érhető el az adott szöveg. Az egyes szövegelemlek betűhű változatát szöveges és PDF formátumban is közreadjuk. Ha az általunk feldolgozott kiadás elérhető az interneten, akkor nem készítettünk hozzá PDF változatot. Ha a szövegnek elkészült a normalizált változata is, azt is elérhetővé tesszük.

Minden szövegelemleket betűhű változata elérhető, de nem mindegyikhez készült normalizálás és morfológiai elemzés. Az alábbi szövegelemleknek van normalizált változata is: Bécsi kódex, Birk-kódex, Bod-kódex, Czech-kódex, Festetics-kódex, Guary-kódex, Jókai-

kódex, Kazinczy-kódex, Könyvecse, Miskolci töredék, Székelyudvarhelyi kódex eleje, továbbá minden rövidebb szövegemlék, kivéve a Csíziót. Az alábbi szövegemlékekhez készült morfológiai elemzés is: Festetics-kódex, Guary-kódex, Jókai-kódex, Könyvecse.

2. A mondat alapszerkezetének változásai

A projektum másik feladata az eddigi magyar nyelvtörténeti kutatások során elhanyagolt mondattani változások vizsgálata volt. Fő célunk az ómagyar mondattan elemzése volt, de számos változást az újmagyar korig végigkísértünk. Amennyire az adatok lehetővé tették, ősmagyar mondattani jelenségeket is rekonstruáltunk.

A mondat alapszerkezetének vizsgálata során megállapítottuk, hogy az ősmagyar mondat SOV mondatszerkezetet örökölt az ugor alapnyelvből. A korai ómagyar hamar visszaszoruló archaizmusai és a mai hanti és manysi összevetésével megállapítottuk, hogy az ősmagyar mondatban összeforrtak a tematikus és az információszerkezeti szerepek: az alany egyszersmind topik is volt, a tárgy pedig vagy fókuszként, vagy másodtopikként funkcionált. Másodtopik szerepét először a tárgyon megjelenő todalék, később az igei megjelenő egyeztető morféma jelölte. Ha a tárgy prominensebb volt a személyhierarchiában mint az alany, a számos nyelvből ismert inverz egyeztetési tilalom megakadályozta másodtopikként való megjelölését. Ennek grammatikalizálódása az 1. és 2. személyű tárgyon hiányzó tárgyra, ill. az 1. és 2. személyű tárggyal való egyeztetés hiánya.

Az ősmagyar SOV alapszórendjének további bizonyítékai a korai ómagyarban még megtalálható, szigorúan SOV szórendű, jelöletlen tárgyat tartalmazó ige-igeneves alárendelő mondatok, az összetett igealakokban megőrződött ige-segédige szórend, és az *-e* kérdő mondatbevezetőnek a korai ómagyarban még megőrzött mondatvégi pozíciója. Az ómagyar SOV mondatszerkezet maradványa a korai ómagyarban uralkodó ige-igeneves alárendelés is.

Mint kimutattuk, első fennmaradt összefüggő nyelvemlékünkre, a Halotti beszéd és könyörgésre már az SVO mondatszerkezet a jellemző. Hipotézist állítottunk fel –a rokon nyelvek most zajló változásait is figyelembe véve – arról, hogy mi váltotta ki az SOV → SVO váltást, és hogyan zajlott le a folyamat.

Megmutattuk, hogy az alapszórend megváltozása a mondattan minden szintjén mélyreható változásokat indított el; megindult a fejező szerkezetek fejeződéjévé való alakulása. A mondat szintjén például a tagadószó az igehez csatolt adverbiumból tagadó egyeztetést indukáló operátorrá változott. A tagadó egyeztetés kialakulása előtti állapotot nemcsak a korai kódexek hanem mai nyelvi kövületek is őrzik (pl. *semmitmondó*, *semmittevés*, *semmirekellő* stb.) Abban a folyamatban, melynek során a *nem* tagadószó

beépült a *semmi*, *senki* stb. névmásokba, majd elveszítette tagadó erejét, s ezáltal szükségessé vált a tagadószó újbóli kitétele, kimutattuk a más nyelvekből is ismert tagadó ciklust. Érdekes változásokon ment át a kérdőszó pozíciója is, például a legkorábbi nyelvemlékek beágyazott kérdéseiben még nem jelent meg a *hog*y kötőszó; a kérdőszó állt a mondatbevezető helyén.

Foglalkoztunk az összetett, időt és aspektust is jelölő ómagyar igealakok kialakulásával és kihalásával. Bizonyítottuk, hogy téves az összetett igealakok kialakulásának latin hatással való magyarázata. Amellett érveltünk, hogy az összetett igei paradigmák ótörök hatásra alakultak ki az ősmagyar korban, és az ugor alapnyelvből hozott személyragos igealakok újraelemzésével és kiterjesztésével jöttel létre. Az összetett igealakok kihalását az igeekötőrendszer kiépülésével hoztuk összefüggésbe. A szituációs aspektus, pontosabban a telikusság igeekötővel való jelölése redundánssá tette a nézőponti aspektus összetett igealakokkal való kifejezését. Abban a folyamatban, melynek során az uráli eredetű *-a/e* múltidőjelet a *-t* befejezettségjel váltotta fel, más nyelvekből is ismert szemantikai újraelemzést mutattunk ki. Foglalkoztunk a tárgyas ragozás kialakulásával, és kétszeri funkcióváltásával (ún. exaptációjával). Amellett érveltünk, hogy a tárgyas ragozás először a tárgy másodtopik szerepét jelölte, majd a tárgy határozottságának jelölésére szolgált. Első funkcióját a topikmozgatás kialakulása tette feleslegessé, második funkcióját a névelőrendszer kifejlődése. Mai szerepe a rejtett tárgyi névmás engedélyezése.

3. A főnévi kifejezés szerkezetének változásai

Számos, az ómagyar főnévi kifejezések belső szerkezeti tulajdonságait érintő kérdést is vizsgáltunk, különös tekintettel a határozott névelő kialakulására, használatának korai jellegzetességeire. Ennek értelmében elemeztük a határozottság szemantikai és szintaktikai jelenségét az ómagyar korban, illetve kimutattuk az ómagyar és a mai magyar főnévi csoport strukturális különbségeit. Az adatgyűjtés és rendszerezés során részben hagyományos filológia módszerekkel dolgoztunk, megvizsgáltuk különböző korokból származó bibliafordítások párhuzamos szövegrészeit is, továbbá gépi keresést is alkalmaztunk, amely egyszersmind hozzájárult az projekt keretein belül épülő Ómagyar korpusz teszteléséhez és javításához. E kutatások eredményeképpen kimutattuk, hogy az ómagyar korban a névelő fokozatosan, több lépcsőben grammatikalizálódott a határozottság nyelvtani kódolására. Ugyanis nemcsak az univerzálisnak tekinthető, szemantikai és pragmatikai értelemben vett határozottság nyelvtani kódolása nyelvspecifikus, hanem az is, hogy e szemantikai mező mekkora részét grammatikalizálja egy nyelv éppen a határozott névelővel. Ez a tartomány idővel módosulhat két nyelvallapot között, például egy adott nyelv története során

előfordulhat, hogy a grammatikalizáció egyre nagyobb tartományban érvényesül, vagyis a névelő használata fokozatosan kiterjed. Ennek értelmében a legkorábbi kéziratos kódexek névelőhasználatának, illetve az ezt követő szinkrón állapotoknak a vizsgálatával megkíséreltük a határozottsági paraméter grammatikai kódolásának fokozatos változását nyomon követni. Az első, biztonsággal rekonstruálható szinkrón nyelvállapotban, a kései ómagyar kor első felében (14. század vége – 15. század első fele) a határozott névelő használata abban tért el a maitól, hogy jóval szűkebb tartományban érvényesült: a korai névelő csak akkor jelenik meg, amikor ténylegesen szükség van a determináltság külön jelölésére, mivel egyéb eszközök nem teszik lehetővé a referenciális azonosítást. A névelő tehát rendszerint elmarad a tulajdonnevek és az inherensen egyedi nevek előtt, továbbá a generikus olvasatú főnévi kifejezésekben, mivel ezek referenciája szemantikai értelemben függetlenül azonosítható. Szükségtelen volt a névelőzés mutató névmási módosítók és birtokos kifejezések jelenlétében, illetve elmaradhatott nem argumentum szerepű összetevők esetében is. Gépi korpuszvizsgálattal viszont kimutattuk a névelőhasználat korszakon belüli arányos növekedését, továbbá azt is, hogy a felsorolt kontextus közül melyikben tapasztalhatunk nagyobb eltolódást, mely szintaktikai környezetekben terjedt számszerűen kimutatható módon a névelő használata.

Részletesen vizsgáltuk a határozott névelő megjelenésének és több környezetben való elterjedésének strukturális következményeit is, így a mutató névmási módosítóval, illetve a birtokos kifejezéssel bővített konstrukciókat. Az itt tapasztalható szórendi változások teljesebb megértése érdekében középmagyar adatokat is figyelembe vettünk. Megállapítottuk, hogy a főnévi frázis bal perifériájának átrendeződése a mai szinkrón állapotig több lépcsőben történt, ahogy ezt a jellegzetes középmagyar szórendi minták tanúsítják.

Új stratégiák jelentek meg ugyanis a mutató névmással történő módosítás kifejezésére, köztük a determináns duplázást alkalmazó mintával (pl. *ez/az a könyv*). A duplázás azonban eleinte lazább szerkezetet eredményezett, így – a mai magyartól eltérően – a mutató névmási módosítót még számos elem választhatta el a névelőtől (kötőszók, igekötők, az *-e* partikula, sőt akár egy ige is), és a birtokos kifejezést is megelőzte, ami a mai magyartól eltérő szórendet eredményezett (pl. *azokat a szerzetesnek könyveit*). Ezt a hagyományosan értelmezős szerkezetnek tekinthető megoldást szintaktikai modellünkben adjunkciónak azaz fráziscsatolásnak ábrázoltuk. A mai magyar mutató névmási módosító azonban szigorúan szomszédos a névelővel, és a birtokos kizárólag előtte jelenhet meg. Ezt a változást úgy modelleztük, hogy a névmási módosító újra a determinánsi frázis szerkezeti tartományába integrálódott – oda, ahol tulajdonképpen eredetileg is volt, mielőtt még a névelő használat

drasztikusan elterjedt volna és az egyértelműsítés igénye létrehozta volna a duplázó szerkezetet. A *-nak/-nek* ragos birtokos kifejezés ezek után már szabályosan a kiterjesztett főnévi frázis legkülső rétegeként jelenik meg. Ez a rekonstrukció megmagyarázza a mutató névmási módosítók kivételes viselkedését a mai magyar főnévi kifejezésekben is: helyzetüket tekintve már szerkezetileg kötöttek ugyan, felépítésükben mégis megőrizték korábbi függetlenségüket és morfológiailag is megjelenítik az egyeztető jegyeket a főnév számával és esetével.

4. Az univerzális és egzisztenciális kvantifikáció változásai

Vizsgáltuk az univerzális kvantifikáció és a disztributivitás ómagyar, korai középmagyar kori eszközeit és a *vala+névmás* összetételű kifejezések vonatkozó névmási értelmezését is. Az univerzális kvantifikáció eszközeit számba véve kimutattuk a *mind* és a *minden* korai és világos elkülönülését, mind a szintaxisuk, mind az értelmezési lehetőségek tekintetében. A *mind*-ről kiderült, hogy nem felel meg a hagyományos logika univerzális kvantorának, elsősorban azért nem, mivel nem kötelezően disztributív (megfér a kollektív predikátumokkal, a kölcsönös névmásokkal és az olyan, par excellence disztributivitást jelölő kifejezésekkel, mint a *ki-ki* vagy a *fejenként*). A *mind* az ómagyar és a középmagyar korban radikálisan alulspecifikált maximalitási operátornak felelhetett meg, hiszen "egész(en)" értelmezéssel módosíthatott elvont főneveket vagy (speciálisan átértelmezett) megszámlálható főneveket (*mind ez világ*, *mind az város* típusú kombinációkban), de egy (időbeli vagy térbeli) skála végpontját is jelölhette.

A *mind* és a *minden* szintaxisának és szemantikájának korai és világos differenciálódása azzal az eredménnyel is járt, hogy a *minden* sokkal inkább emlékeztet a klasszikus logika univerzális kvantorára: disztributív, és emellett szintaktikailag determinált az, hogy a mondatból mi lesz Restriktor és mi lesz Szkópusz. A *minden* sokkal inkább alkalmas volt általános ítéletek kifejezésére, mint a *mind* (hiszen a *mind* elsősorban határozott, a diskurzusból ismert kifejezésekhez társult).

A mai magyarban egzisztenciális értelmezésű *vala+névmás* kombináció az ómagyarban és a középmagyarban lehetett vonatkozó névmás (amikor is egyfajta univerzális/maximális értelmezést lehetett neki tulajdonítani), és lehetett korlátlan választékot kifejező elem (a mai *akármi (is)*, *bármilyen (is)*-hez hasonlóan). Emellett a *valaki*, *valami* nem vett fel kötelezően magas hatókört az azonos tagmondati tagadáshoz viszonyítva (szemben a mai *valaki*-vel, amely ezek szerint pozitív polaritású elem). A fejezetben megfogalmazott hipotézis szerint a *vala+névmás* kifejezés vonatkozó mellékmondatok specifikálójában

állhatott elő, azáltal, hogy a főmondatbeli *vala* ige átértelmeződött, azaz, a beszélők megváltozott grammatikája szerint a mellékmondat CP rétegéhez tartozott, és egy összetevőt alkotott a közvetlenül utána jövő névmással. Ha a *vala*- "operátort" az angol *any*-hez vagy *ever*-hez hasonló módon kezeljük, ezzel magyarázatot tudunk adni a *valami* ómagyar kori 'korlátlan választék' értelmezésére, és arra, hogy miért nem volt kötelezően tág hatóköre az azonos tagmondati tagadáshoz viszonyítva. Javaslatunk szerint az ómagyar kori *valaki* határozatlan (egzisztenciális) kifejezés volt, szintaktikai környezettől függetlenül. Ebből elvben belátható, miért lehet a mai *valami* tisztán egzisztenciális.

6. Az igeneves kifejezések története

Az igenevek korábbi hagyományos nyelvészeti keretben készült vizsgálatai (pl. Károly 1956) morfológiai alapúak voltak. Ebből fakadóan a csoportosítás alapja az igenévi végződés fonológiai formája volt: például minden fajta *-t* melléknévi igenevet (*a leomlott fal, az anyám sütötte kenyér, a keze aszott ember, Mária ülte fű*) egy viszonylag homogén csoportnak tételeztek; holott ezeknek az igeneveknek mások a képzési feltételei, és eltérő tulajdonságokkal rendelkeznek az egyeztetés és a bővítményfelvevő képesség tekintetében is. A korábbi szerkezet alapú tárgyalások (pl. Laczkó 2000) egyáltalán nem foglalkoztak azokkal az igenevekkel, amelyek az ómagyarban még megtalálhatóak de a mai magyarban már nincsenek meg. A projektum eredményeit bemutató magyar kötetben az összes, a mai magyarban és az ómagyarban megtalálható igenevet szerkezeti alapon tárgyaltuk: a mondattanilag különböző, de azonos igenévi végződéssel rendelkező igeneveket szétválasztottuk, és az összes igenévtípus mögöttes reprezentációját megadtuk. Ez a különböző igenévtípusok eddigi legátfogóbb, szerkezet alapú tárgyalását eredményezte. Mivel szerkezeti alapon először tárgyaltuk a csak az ómagyarban meglevő igenévtípusokat, ezen típusokra új terminológiát is be kellett vezetni.

A magyar nyelvtörténeti munkák ez idáig arra fókuszáltak, hogy az egyes igenévtípusok változásait egyenként leírják, de nem tárgyalták a rendszerszinten bekövetkező változásokat: azaz azokat a változásokat, melyek egyszerre több igenévtípust is ugyan úgy érintettek. A projekum által azonosított nagy ívű, rendszerszintű változások a következők. I) Az igenevek fokozatosan kiszorultak a nyelvből, a véges alárendelés az igenevek rovására terjedt. Ez abban nyilvánult meg, hogy szinte az összes igenév disztribúciója szűkült, több típus vesztett produktivitásából, egyes igenevek pedig teljesen eltűntek. II) A megmaradt igenevek egyre több tulajdonságban kezdtek függni a főmondattól, azaz egyre tipikusabb igenévi kategóriákká váltak. Ez abban nyilvánult meg, hogy több igenévtípusban kötelezővé

vált a kiejtetlen, koreferens alany a korábbi kiejtett, független referenciájú alany helyett, és több igenévtípus elvesztette egyezési képességét az alannal. A projektum eredményeit összefoglaló angol nyelvű kötetben az igenevek történetét e rendszer szintű változások mentén mutattuk be.

A projektum azon hipotézisét, hogy az ősmagyar nyelv SOV szórendű, fejkéű nyelv volt, az igenevek alapján kétféle bizonyítékkal támogattuk. Azt javasoltuk, hogy a szigorúan fejkéű igenevek (pl. az *anyám süttötte kenyér* és a keze aszott *ember* típus) az ősmagyar igenkéű szórend maradványa az ómagyarban, továbbá az ómagyarban még megtalálható, mindig szigorúan preverbális jelöletlen tárgyak is ilyen ősi maradványok.

A korábbi nyelvtörténeti munkák megemlítették az anti-egyző, azaz az alany személytől és számától függetlenül E/3 egyeztetést viselő infinitívuszok (pl. *nem jöttem hínia az igazakat*) létezését, de nem vizsgálták, hogy az anti-egyzés milyen feltételek között jöhet létre. Kimutattuk, hogy az anti egyzés a kontrolllos infinitívuszok körében opcionális jelenség, és a minimalista elmélet eszköztárát felhasználva modelláltuk az anti-egyző infinitívuszok szintaktikai szerkezetét. Az elemzés megerősíti azt a nemzetközi irodalomban nemrégiben felmerült elképzelést, hogy az erős/gyenge fázis megkülönböztetés nem csak az igei csoport (vP), hanem a mondat (CP) szintjén is releváns, és az infinitívuszok erős és gyenge fázisok is lehetnek. Az ehhez kapcsolódó eredményeket a *Nyelvtudományi Közlemények*-ben publikáltuk. Várható továbbá egy könyvfejezet egy angol nyelvű tanulmánykötetben a Mouton kiadónál.

Új generatív elemzést adtunk az ómagyar –t gerundium (pl. *jártomban, keltemben*) szerkezetére. Az új elemzés először próbál választ adni arra kérdésre, hogy ezek az igenevek miért viselnek kötelezően birtokos morfológiát. Az eredményeket az Olomouc Linguistics Colloquium és 11th International Conference on the Structure of Hungarian konferenciákon ismertettük, és az utóbbi konferencia konferenciakötetében tervezzük megjelentetni.

A kizárólag a korai kódexekben megtalálható –atta/ette igenevekről a korábbi leíró igényű szakirodalom (pl. Károly 1956) azt tartotta, hogy kétféle alanyt engednek meg: önálló referenciával rendelkező és kiejtetlen, koreferens alanyt. A koreferens alany referenciáját a főmondat alanya vagy tárgya határozza meg. Kimutattuk, hogy egyes esetekben a koreferens alany referenciáját a főmondat részes esetű (-nak/nek ragos) összetevője, nem pedig az alany vagy a tárgy határozza meg.

6. A névutós kifejezés változásai

Vizsgáltuk a névutók grammatikalizálódását és a névutós kifejezésen belüli szerkezeti változásokat. Elemeztük a birtokos szerkezetben keletkező névutókat, és azt állapítottuk meg, hogy az ómagyar korban ezek közül több is olyan szerkezeti variációban vesz részt, amely megkérdőjelezi névutói jellegüket. Az utóbbi évek összehasonlító vizsgálatainak eredményeit felhasználva azt mondtuk, hogy ezek az elemek nem névutók (nem a PP-n belüli P fejben vannak generálva), hanem egy átmeneti kategória fejét töltik ki (AxPartP a szerkezetben), ez teszi lehetővé, hogy birtokoshoz hasonló *-nak/-nek* ragos vonzattal és egyeztető morfémat viselve jelenjenek meg a szerkezetben.

Az ún. határozói eredetű névutók kialakulását kétlépcsős grammatikalizációs folyamattal jellemeztük: első lépésben kialakul egy névutó, majd ez a névutói egy értelmezői szerkezeten keresztül újra beépül egy PP-be, és ott funkcionális adpozícióvá válik, amit szerkezetileg egy funkcionális fej (p) reprezentál. Ezek a névutók vagy valamilyen más PP mellett jelennek meg (erre utal a „ragvonzó névutó” elnevezés), vagy egy implicit deiktikus vonzattal. Ez utóbbi esetben a mondatban önállóan viselkednek helyhatározóként, de szerkezetileg mindig pP-t alkotnak. Az ómagyar korban ez a csoport még kevesebb elemmel rendelkezik, mint ma, de a középmagyar korra nő az elemek száma. Szerkezetileg funkcionális fejként (p-ként) elemeztük az igekötőket is, amelyek szintén megjelenhetnek PP vonzattal is a mondatban. A legrégebbi igekötők már az ómagyar korban is néha pusztán funkcionális elemként viselkednek, azaz nincs mindig helyhatározói jelentésük (főként a meg-re jellemző ez). Az írásos korszak során nő az igekötők száma, és használatuk is egyre gyakoribbá válik.

7. Az alárendelő összetett mondatok változásai

A vizsgálat kiterjedt a véges alárendelő mondatok történeti változásaira is, különös tekintettel a hasonlító mellékmondatokra. Ennek során megmutattuk, hogy a magyar alárendelő mondatokban bekövetkezett változások olyan általános elvek mentén írhatók le, amelyek értelemszerűen más nyelvekben is megfigyelhetők, ugyanakkor a magyar szerkezeteknek vannak olyan sajátosságai, amelyek más nyelvekben nem vagy csak sporadikusan adathozhatóak, és éppen ezért a magyar nyelvvel kapcsolatos eredmények a nyelvelmélet számára is különösen fontosak lehetnek.

A hasonlító szerkezetekben megfigyelhető változások korábban is ismertek voltak, azonban formális, generatív szempontú leírásukra elsőként most került sor. A hasonlító mellékmondatot eredetileg a *hogy* kötőszó vezette be, vagyis nem volt a szerkezetre speciálisan jellemző mondatbevezető; egyenlőtlenséget kifejező hasonlításban pedig a

szerkezet a *nem* elemet is tartalmazta, amely a negatív polaritás jelölője. Később a *mint* is megjelent, a *hogynem mint* szekvenciát eredményezve, végső soron pedig mind a *nem*, mind pedig a *hogynem* eltűnt a szerkezetből, így a mai magyarra már csak a *mint* jellemző. A kutatásunk megmutatta, hogy nem pusztán funkciócsereről van szó, hanem a *mint* megjelenése és újraértelmezése, végül pedig a *hogynem* eltűnése szigorú grammatikalizációs szabályok alapján írható le. Ennek értelmében a *mint* először vonatkozó névmásként jelent meg, majd mondatbevezetővé grammatikalizálódott, végül pedig már egyedül is képes volt jelölni nemcsak a hasonlító mondat típust, hanem a véges alárendelést is.

A *mint* újraértelmezése és grammatikalizációja jelentős részben a relatív ciklus folyamatával írható le, amely számos nyelvben megfigyelhető. A magyarban rámutattunk, hogy a mai alárendelő kötőszók (*hogynem ha mint mert*) mindegyike ezen a többlépcsős változási soron ment keresztül, ugyanakkor egymástól eltérő időben. Ennek következtében az ómagyar és a középmagyar kor során a mondatbevezetők kombinálódhattak is egymással, hiszen amennyiben más szintaktikai pozíciót foglaltak el, úgy nem volt akadálya annak, hogy egyszerre jelen legyenek a szerkezetben. Ily módon a magyarban a kombinációknak igen gazdag rendszere jött létre; a kombinációk egy része két mondat típus kereszteződése (pl. *ha + mint*), más kombinációkban pedig két, részben azonos funkciójú elem állt együtt (pl. *hogynem mint*), megint más kombinációkban pedig a *hogynem* kötőszó általános véges alárendelő kötőszóként kombinálódik egy, az adott mondat típusra specifikusan jellemző kötőszóval (pl. *hogynem mert*). A rendszer a kettős kötőszókon kívül kiegészül még a polaritást jelölő tagadó elemekkel, illetve a kötőszóknak a vonatkozó névmásokkal való kombinációi is ide sorolhatók. Míg más nyelvekben rendszerint nem jelentkezik mind a három típus, kutatásaink világosan megmutatták, hogy a magyar nyelv történetének egy adott szakaszában (az ómagyar korban, illetve a korai középmagyarban) ezek a kombinációk mind produktívak voltak, ugyanakkor a kombinációk megjelenése szabályokkal leírható módon történt.

Megvizsgáltuk azt is, hogy az egyes kötőszók kombinációinak hogyan és milyen feltételek mellett jöhetett létre az inverz párja (tehát például a *hamint* kombinációnak a *mintha*), és a kutatás arra az eredményre vezetett, hogy ezekben az esetekben egy alsóbb pozícióban generálódó kötőszó felmozgott a felsőbb fejpozícióba, és balról adjungálódott az ott lévő kötőszóhoz. Ezek tehát morfológiai komplex fejek, amelyek később így grammatikalizálódtak, vagyis szintaktikailag nem szétbontható egységgé váltak. Az ilyen típusú alakulatok létrejötte szintén nem csak a magyar nyelvben figyelhető meg, ugyanakkor az, hogy ez rendszerszerűen, lényegében bármely kötőszói kombináció esetében létrejöhetett, elsősorban a magyar nyelv morfológiai rugalmasságának köszönhető. Ez a típusú változás az

egyszerű kötőszók és a kötőszói kombinációk rendszerének egyébként is megfigyelhető gazdagságával együtt az ómagyar korban és a középmagyar kor elején rendkívül színes képet eredményezett a véges alárendelő mellékmondatok bal periferiáján. Ez a sokszínűség már a korábbi szakirodalomban is ismert volt, ugyanakkor a sokszínűség mögötti szabályrendszer feltérképezése a mostani kutatás eredménye.

Az egyes kombinációk kialakulása mellett azzal a kérdéssel is foglalkoztunk, hogy mely alakulatok maradtak meg a magyar nyelvben, és melyek haltak ki; arra az eredményre jutottunk, hogy a mai magyarban a többszörös kötőszók már nincsenek meg, viszont a morfológiailag komplex, de grammatikalizálódott fejek megmaradtak. Vagyis a kötőszói kombinációk rendszerének nemcsak a kialakulása és továbbélése, hanem a részbeni eltűnése is szabályokkal jól leírható jelenségcsoport.

A véges alárendelő mellékmondatok funkcionális bal periferiáján kimutatható változások jelentős része az ómagyarra, illetve a középmagyarra tehető; a négy fő kötőszó természetesen már korábban is létezett a nyelvben, azonban nem mindegyikük grammatikalizálódott kötőszóvá az ómagyar kor előtt. Ennek oka leginkább az, hogy a magyarban korábban jóval nagyobb szerepet töltött be a nemvéges alárendelés, és így a véges alárendelésre jellemző grammatikalizációs folyamatok értelemszerűen leginkább ahhoz az időszakhoz köthetők, amikor a véges alárendelés szerepe (az SOV szórendet felváltó SVO szórend terjedésével párhuzamosan) megnőtt. Egyrészt a véges alárendelő mondatok gyakorisága is kimutathatóan megnőtt (ez jól látszik a párhuzamos bibliafordítások összevetéséből), másrészt egyre összetettebb rendszer alakult ki azáltal, hogy újabb elemek grammatikalizálódtak a bal periferián, illetve kombinálódtak egymással.

8. Disszemináció

A projektum végeredménye egy magyar nyelvű kötet: *Magyar történeti mondattan*, megjelenés alatt az Akadémiai Kiadónál, valamint egy angol nyelvű kötet: *Hungarian Diachronic Syntax*, megjelenés alatt az Oxford University Pressnél. Kutatási eredményeinket 28 angol nyelvű nemzetközi publikációban: folyóiratcikkben és könyvfejezetben, valamint 24 magyar nyelvű folyóiratcikkben, ill. könyvfejezetben is közzétettük. Nemzetközi konferenciaelőadásaink száma 32, hazai konferenciaelőadásaink száma 39.